Supplemental data

TOWARDS CLASSIFICATION OF BRCA1 MISSENSE VARIANTS USING A BIOPHYSICAL APPROACH

Pamela J.E. Rowling, Rebecca Cook & Laura S. Itzhaki*

Prediction methods

A number of prediction algorithms were used in this study. SIFT (1) and I-Mutant 2.0 (2) do not require the structure of the domain to be known. For the algorithms that do require a structure the pdb file 1JNX was used (3). SIFT (http:/blocks.fhrc.org/sift/SIFT.html (1)) scores the change at a particular position according to the conservation at this site and the type of amino acid change. I-Mutant 2.0 (http://gpcr2.biocomp.unibo.it/~emidio/I-Mutant/I-Mutant.htm) is a neural network-based prediction method that has been trained on a data set of known mutations (2). This method can use either the protein sequence or a structure for the basis of the prediction; we used the structure. Site Directed Mutator (SDM) algorithm uses a set of conformationally-constrained environment-specific substitution tables to calculate differences in stability scores between the folded and unfolded states (4). PolyPhen (www.bork.embl-heidelberg.de/PolyPhen (5)) uses multiple sequence alignments and adds structure-based criteria to predict the effect of mutations. D-Fire (http://sparks.informatics.iupui.edu/hzhou/mutation.html) calculates the atomic stability of the structure with the substitution at a specific position (6,7). Fold-X (http://foldx.crg.es/) calculates the free energy of the protein based on its structure (7). The ERIS server (http://troll.med.unc.edu/eris/) calculates the change in protein stability induced by mutations. The program has not been trained on a dataset and uniquely can model backbone flexibility to allow incorporation of a mutation (8).

Figure Legend

Fig. S1. Temperature dependence of the free energy of unfolding of wild-type BRCA1 BRCT. The free energy change for each of the two unfolding transitions is plotted as a function of temperature for: the transition between the folded state and the intermediate (circles) and between the intermediate and the unfolded state (diamonds). The data were fitted using equation 2. For the transition between the folded state and the intermediate, the following parameters were obtained from the fit: $\Delta H_{I\text{-}F}$ = 145 kcal mol^{-1}, $\Delta C_{p(I\text{-}F)}$ = 4.3 kcal K^{-1} mol^{-1} and $T_{m(I\text{-}F)}$ = 39 ^{o}C. For the transition between the intermediate and the unfolded state, the parameters obtained were: $\Delta H_{U\text{-}I}$ = 60 kcal mol^{-1}, $\Delta C_{p(U\text{-}I)}$ = 1.8 kcal K^{-1} mol^{-1} and $T_{m(U\text{-}I)}$ = 48 ^{o}C.

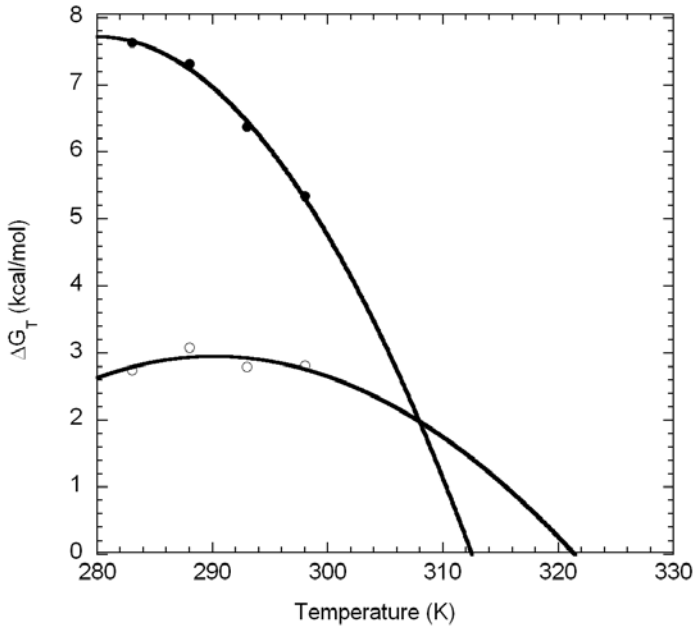

Table S1. Comparison of the experimental and predicted effects of mutation on the thermodynamic stability, $\Delta\Delta G_{U\text{-}F}$, of BRCA1 BRCT domains. The variants were grouped according to their experimentally-determined effects on stability. A positive value of $\Delta\Delta G_{U\text{-}F}$ indicates that the mutation is destabilising and a negative value stabilising. The standard errors on $\Delta\Delta G_{U\text{-}F}$ were 5-10%. * denotes mutant proteins that could not be characterised because they expressed in inclusion bodies (IB).

NOT DESTABILISING

Variant	Experimental $\Delta\Delta G_{U\text{-}F}$ (kcal mol^{-1})	FoldX $\Delta\Delta G_{U\text{-}F}$ (kcal mol^{-1})	ERIS $\Delta\Delta G_{U\text{-}F}$ (kcal mol^{-1})	SDM $\Delta\Delta G_{U\text{-}F}$ (kcal mol^{-1})	DFIRE $\Delta\Delta G_{U\text{-}F}$ (kcal mol^{-1})	I-Mutant 2.0 $\Delta\Delta G_{U\text{-}F}$ (kcal mol^{-1})	PolyPhen Effect	SIFT Classification
M1663L	-0.51	0.06	2.10	0.83	-0.54	-0.15	probably damaging	neutral
M1663K	-0.03	0.17	4.40	0.23	0.29	-1.57	probably damaging	neutral
A1669S	-0.16	0.65	2.13	3.05	0.65	-0.58	benign	neutral
R1699L	-0.99	-0.20	0.17	-2.03	-1.17	-0.10	probably damaging	deleterious
R1699Q	-1.83	1.34	0.90	1.30	-2.53	-0.62	probably damaging	deleterious
C1787S	-0.37	0.70	1.12	5.96	2.17	-1.16	probably damaging	deleterious
P1806A	0.06	2.12	-0.33	-2.56	0.06	-0.13	benign	neutral

MILDLY DESTABILISING

Variant	Experimental $\Delta\Delta G_{U\text{-}F}$ (kcal mol^{-1})	FoldX $\Delta\Delta G_{U\text{-}F}$ (kcal mol^{-1})	ERIS $\Delta\Delta G_{U\text{-}F}$ (kcal mol^{-1})	SDM $\Delta\Delta G_{U\text{-}F}$ (kcal mol^{-1})	DFIRE $\Delta\Delta G_{U\text{-}F}$ (kcal mol^{-1})	I-Mutant 2.0 $\Delta\Delta G_{U\text{-}F}$ (kcal mol^{-1})	PolyPhen Effect	SIFT Classification
M1652I	1.11	1.93	2.91	-0.45	-0.98	-0.69	probably damaging	neutral
L1664P	1.18	1.85	14.86	2.46	3.08	-1.48	probably damaging	deleterious
V1665M	2.22	8.59	6.16	0.43	-0.22	-1.27	benign	deleterious
D1692N	1.15	0.07	-0.71	0.39	-0.19	-0.46	possibly damaging	deleterious
G1706A	1.12	6.43	-0.61	-2.54	-1.42	-1.92	benign	deleterious
R1751Q	1.57	-1.38	0.10		-0.66	-1.03	possibly damaging	unclassified
T1773S	0.48	-1.61	0.00	0.42	0.12	-0.09	benign	deleterious
D1778N	0.76	0.21	0.11	0.62	-0.24	0.28	possibly damaging	neutral
G1788V	1.84	39.29	14.35	0.86	-0.12	0.01	probably damaging	deleterious

MODERATELY DESTABILISING

Variant	Experimental $\Delta\Delta G_{U-F}$ (kcal mol^{-1})	FoldX $\Delta\Delta G_{U-F}$ (kcal mol^{-1})	ERIS $\Delta\Delta G_{U-F}$ (kcal mol^{-1})	SDM $\Delta\Delta G_{U-F}$ (kcal mol^{-1})	DFIRE $\Delta\Delta G_{U-F}$ (kcal mol^{-1})	I-Mutant 2.0 $\Delta\Delta G_{U-F}$ (kcal mol^{-1})	PolyPhen Effect	SIFT Classification
R1699W	2.30	3.71	14.16	-2.10	0.89	-0.77	probably damaging	deleterious
V1736A	4.20	0.02	6.06	1.16	1.71	-2.62	benign	deleterious
M1783T	3.73	4.41	5.53	3.75	2.26	-0.92	probably damaging	deleterious
G1788D	3.97	14.62	3.68	2.07	0.08	-0.63	possibly damaging	deleterious
V1808A	2.40	3.44	6.06	2.71	2.55	-1.01	benign	unclassified
A1843P	4.89	8.43	18.12	4.48	1.10	-0.71	probably damaging	neutral

VERY DESTABILISING

Variant	Experimental $\Delta\Delta G_{U-F}$ (kcal mol^{-1})	FoldX $\Delta\Delta G_{U-F}$ (kcal mol^{-1})	ERIS $\Delta\Delta G_{U-F}$ (kcal mol^{-1})	SDM $\Delta\Delta G_{U-F}$ (kcal mol^{-1})	DFIRE $\Delta\Delta G_{U-F}$ (kcal mol^{-1})	I-Mutant 2.0 $\Delta\Delta G_{U-F}$ (kcal mol^{-1})	PolyPhen Effect	SIFT Classification
D1692Y	IB*	-0.36	0.85	1.22	-1.90	1.39	probably damaging	deleterious
A1708E	IB*	16.55	17.93	2.69	2.39	-0.72	probably damaging	deleterious
S1715C	IB*	1.85	5.19	6.96	-1.96	-1.93	benign	deleterious
S1715N	IB*	13.35	7.30	1.46	0.75	-0.35	probably damaging	deleterious
S1715R	IB*	16.41	10.30	0.39	1.85	-0.96	probably damaging	deleterious
L1764P	IB*	7.32	3.73	3.25	3.31	-1.87	probably damaging	unclassified
I1766S	IB*	4.44	11.34	5.06	4.41	-1.98	probably damaging	deleterious
M1775R	IB*	4.96	6.13	2.43	1.07	-0.96	probably damaging	deleterious
L1780P	IB*	5.77	18.58	4.68	5.08	-1.50	probably damaging	deleterious
V1833M	IB*	6.28	7.24	0.65	0.55	-0.94	benign	unclassified
W1837G	IB*	7.09	11.90	6.29	7.05	-3.60	probably damaging	deleterious
W1837R	IB*	4.68	10.42	2.45	5.37	-1.88	probably damaging	deleterious
S1841N	IB*	12.67	12.10	1.03	0.43	-0.04	probably damaging	deleterious
Y1853C	6.04	3.45	5.14	6.16	1.11	0.28	probably damaging	deleterious

References for Supplemental Data

1. Ng, P. C., and Henikoff, S. (2003) *Nucleic Acids Res* 31, 3812-3814
2. Capriotti, E., Fariselli, P., and Casadio, R. (2005) *Nucleic Acids Res* 33, W306-310
3. Williams, R. S., Green, R., and Glover, J. N. (2001) *Nat Struct Biol* 8, 838-842
4. Worth, C. L., Bickerton, G. R., Schreyer, A., Forman, J. R., Cheng, T. M., Lee, S., Gong, S., Burke, D. F., and Blundell, T. L. (2007) *J Bioinform Comput Biol* 5, 1297-1318

5. Ramensky, V., Bork, P., and Sunyaev, S. (2002) *Nucleic Acids Res* 30, 3894-3900
6. Zhou, H., and Zhou, Y. (2002) *Protein Sci* 11, 2714-2726
7. Guerois, R., Nielsen, J. E., and Serrano, L. (2002) *J Mol Biol* 320, 369-387
8. Yin, S., Ding, F., and Dokholyan, N. V. (2007) *Nat Methods* 4, 466-467